Ancilla-Assisted Enhancement of Channel Estimation for Low-Noise Parameters

Masahiro Hotta,[1, *] Tokishiro Karasawa,[2, †] and Masanao Ozawa[2, ‡]

[1]*Department of Physics, Faculty of Science, Tôhoku University, Aoba-ku, Sendai, 980-8578, Japan*
[2]*Graduate School of Information Sciences, Tôhoku University, Aoba-ku, Sendai, 980-8579, Japan*
(Dated: Revised August 29, 2005)

In order to make a unified treatment for estimation problems of a very small noise or a very weak signal in a quantum process, we introduce the notion of a low-noise quantum channel with one noise parameter. It is known in several examples that prior entanglement together with nonlocal output measurement improves the performance of the channel estimation. In this paper, we study this "ancilla-assisted enhancement" for estimation of the noise parameter in a general low-noise channel. For channels on two level systems we prove that the enhancement factor, the ratio of the Fisher information of the ancilla-assisted estimation to that of the original one, is always upper bounded by 3/2. Some conditions for the attainability are also given with illustrative examples.

PACS numbers: 03.67.-a, 03.67.Hk, 03.67.Lx, 03.65.Ta

I. INTRODUCTION

One of the formidable obstacles for the realization of quantum computers is decoherence caused by the coupling between computational qubits and the environment. Recent study of quantum error correction has shown that fault-tolerant quantum computing is in principle possible, but it requires that the noise caused by the decoherence should be lower than the very stringent threshold. Obviously, such a statement has a physical meaning only if we have an efficient method for quantitatively estimating very small noise in quantum devices in real experiments. However, if the noise is very small, so is our success probability of observing the disturbance caused by that noise. This difficulty makes evident the demand for the study of optimal quantum estimation of very small nose in general quantum channels based on well-established quantum estimation theory.

Quantum estimation theory was instituted by Helstrom in the late 1960's and has been developed with various applications until recently; for standard reviews we refer to Helstrom [1] and Holevo [2], and see also Hayashi [3] for recent progress. A typical problem of quantum estimation is to ask what is the best observable, possibly in an extended system with ancilla, to measure in order to estimate the true value of θ provided that the system is known to be in one of the state in a given family $\{\rho_\theta\}$. A well-established solution for this problem is given as follows. We call an observable A a (locally) unbiased estimator at $\theta = \theta_0$ if the expectation value $E_\theta[A]$ of A in the state ρ_θ satisfies

$$E_{\theta_0}[A] = \theta_0, \tag{1}$$
$$\partial_\theta E_\theta[A]|_{\theta=\theta_0} = 1. \tag{2}$$

In general there are many unbiased estimators. In order

to select a good one, we consider the variance $V_\theta[A]$ of an arbitrary unbiased estimator A in the state ρ_θ. Then, the quantum Cramér-Rao inequality

$$V_\theta[A] \geq \frac{1}{J(\rho_\theta)} \tag{3}$$

holds for any unbiased estimator A at θ, where

$$J(\rho_\theta) = \text{Tr}[\rho_\theta L_\theta^2] \tag{4}$$

is the (quantum) Fisher information defined through the symmetric logarithmic derivative (SLD) L_θ that is characterized by the relations

$$\partial_\theta \rho_\theta = \frac{1}{2}(L_\theta \rho_\theta + \rho_\theta L_\theta), \tag{5}$$
$$L_\theta^\dagger = L_\theta. \tag{6}$$

The SLD is determined uniquely on the range of ρ_θ, i.e., $L_\theta \rho_\theta = L_\theta' \rho_\theta$ holds for any two SLDs L_θ and L_θ'. The Cramér-Rao inequality (3) follows from a simple application of the Schwarz inequality for the Hilbert-Schmidt inner product. From the equality condition for that the lower bound J_θ^{-1} in Eq. (3) is always achieved by any observable A satisfying

$$A\rho_\theta = (J_\theta^{-1}L_\theta + \theta)\rho_\theta, \tag{7}$$

see Refs. [1, 2] and for a straightforward derivation see Appendix of Ref. [4]. In general, to find an optimal estimator for the true value θ needs prior information on the value θ, which might be collected by prior estimations assuming prior probabilities on the unknown parameter, so that the optimal estimator is considered as an ultimate limit allowed by physics. However, there are some cases in which the optimal estimator can be chosen uniformly over unknown values of θ [5]. In these cases the ultimate limit can be certainly achieved without prior information.

From the quantum estimation theory for state parameters mentioned above, we can construct an estimation theory for unknown parameters of physical processes,

such as coupling constants of the interaction. Suppose that we prepare a quantum system in an initial state ρ_{in} and leave it in an evolution process characterized by an unknown parameter θ. Then, the final state $\rho_{out}(\theta)$ of this process depends on the parameter θ. The problem of finding the optimal estimation of the parameter θ is solved by maximizing the Fisher information J_θ over all the possible initial states ρ_{in} and all the possible observable A in the final state [6, 7]. The above physical process can be represented by a mapping Γ_θ that transform the initial state ρ_{in} to the final state ρ_{out} as

$$\rho_{out} = \Gamma_\theta[\rho_{in}]. \tag{8}$$

It is now fairly well-known that every general state change, called a quantum operation or a quantum channel, such as Γ_θ, physically realizable with probability one should be a trace-preserving completely positive (TPCP) mapping, and conversely that every TPCP map can be realized as a unitary process of the system augmented by an ancilla prepared in a fixed state as shown by Kraus [8, 9]; see also Ref. [10, 11] for the generalization of the above statement to generalized measurements and see Ref. [12] for the latest elaboration.

As pointed out in Ref. [13], one can improve the parameter estimation if a correlation, or in particular an entanglement, is allowed between the input system S and an ancilla A. It should be stressed that in doing so one needs no physical process to occur on the ancilla system A while the system S passes through the channel Γ_θ. In this case, the extended channel is represented as $\Gamma_\theta \otimes id_A$, where id_A stands for the identity channel for A. Then, the improvement can be achieved by the initial preparation of the composite system in an entangled state together with the measurement of the composite system after the process.

Recent progress has been reported on problems for special families of quantum channels, in particular, SU(2) channel [14], a generalized Pauli channel [15], a generalized amplitude-damping channel [16], U(N) channel and its Abelian subgroup channel [17, 18]. A review by Fujiwara [19] is also available. For earlier contributions see also [20, 21, 22, 23, 24, 25]. In this paper, we are devoted to the ancilla-assisted enhancement of Fisher information derived by the quantum Cramér-Rao bound, whereas ancilla-assisted enhancements have been recently investigated within the Bayesian approach [26, 27, 28] and the minimax approach [29].

This enhancement effect not only projects a theoretical profundity of quantum mechanics, but also suggests many physical applications including the low-noise estimation in quantum computing, where the enhanced noise estimation is expected to contribute to developing the quantum error correction and quantum noise reduction technology [30].

We can find another application of the low-noise estimation in elementary particle physics. Today, because of technological difficulties of high-energy experiments, direct researches of new physics far beyond the TeV energy scale are almost impossible. This is one of reasons why the low-energy rare processes predicted by the new physics recently attract much attention. (The CPT symmetry violation in the $K - \bar{K}$ oscillation is one of the typical processes [31, 32, 33, 34].) Clearly, the number of signals for the new-physics evidence is predicted very small, even if the process really exists in nature. The new-physics data should be separated from an enormous number of ordinary data explained by the standard model. This means that the new-physics data can be regarded as a sort of background low noise in the standard data. Hence, we can treat the rare process as a low-noise channel. It is very significant to estimate the intensity of the low noise because indirect information about physics beyond the standard model is obtained. In the estimation, the above ancilla-assisted enhancement may effectively reduce the trial number of the experiment.

In this paper, we study the estimation theory of the parameter characterizing a small noise in a general quantum channel on a system with finite dimensional state space. We can always decompose the quantum channel into two channels so that the input state of the original channel passes through the first noiseless channel and consecutively passes through the second noisy quantum channel called the noise channel. Thus, we can concentrate our attention on the noise channel. We are interested in the case where the noise is so small that the noise channel deviates only a little from the identity channel. In such a case, the channel is called a low-noise channel, and the parameter representing the noise is called the low-noise parameter denoted by ϵ. Let Γ_ϵ be a low-noise channel with low-noise parameter ϵ. We assume that the low-noise parameter is scaled so that Γ_0 is the identity channel. We can formulate natural mathematical requirements for the behavior of the low-noise parameter in a neighborhood of $\epsilon = 0$. It is an interesting problem to figure out how much ancilla-assisted enhancement can be achievable in the estimation of the low-noise parameter ϵ. In this paper we shall discuss this problem and obtain several upper bounds for this ancilla-assisted enhancement factor in the low-noise parameter estimation.

In Section 2, we explain a theorem [13] states that the Fisher information is attained in a pure initial state, so that we can always assume that the input of the channel is a pure state. In Section 3, we discuss parameter estimation for unitary channels, which do not couple with the environment, and show that we have no ancilla-assisted enhancement. Thus, the ancilla-assisted enhancement is possible only for channels coupled with the environment. In Section 4, we introduce the notion of low-noise channels mentioned above with rigorous mathematical requirements, and we obtain a general formula for the upper bound for the ancilla-assisted enhancement factor. In Section 5, we introduce two physical examples of low-noise channel. In Section 6, we give a concrete evaluation of the enhancement factor in two level systems. Let us consider a low-noise channel Γ_ϵ with low-noise parameter ϵ in a two level system S_2. We obtain a universal upper

bound for the enhancement factor η defined by

$$\eta = \frac{\mathcal{L}[\max[J_{S_2+A}]_{\rho_{S_2+A}}]}{\mathcal{L}[\max[J_{S_2}]_{\rho_{S_2}}]} \tag{9}$$

for any finite level ancilla A. Here, ρ_{S_2} is the input in the system S, J_{S_2} is the Fisher information of $\Gamma_\epsilon[\rho_{S_2}]$, ρ_{S_2+A} is the channel input in the composite system $S_2 + A$, J_{S_2+A} is the Fisher information of the output states $(\Gamma_\epsilon \otimes id_A)[\rho_{S_2+A}]$, and $\max[\cdot]_\rho$ stands for the maximum over all the state ρ. As shown later, J_{S_2} and J_{S_2+A} shows a singular behavior $\propto 1/\epsilon$ in the ϵ expansion, and $\mathcal{L}[J]$ is coefficient of $\propto 1/\epsilon$, i.e.,

$$J(\epsilon) = \frac{1}{\epsilon}\mathcal{L}[J] + O(\epsilon^0). \tag{10}$$

The universal upper bound of the enhancement factor η for all the two level systems is given by

$$\eta \leq \frac{3}{2}. \tag{11}$$

This upper bound is attainable by various channels Γ_ϵ, and the corresponding optimal input state is a maximal entangled state, and holds for any low-noise channels on two level systems.

II. THE MAXIMUM IS ATTAINED BY A PURE INPUT STATE: THE FUJIWARA THEOREM

In this section we briefly review an important theorem due to Fujiwara [13]: *the maximum of the Fisher information of output states $\rho_\theta (= \Gamma_\theta[\rho])$ over all possible input states ρ is attained by a pure input state for an arbitrary fixed channel Γ_θ.*

To show this following Fujiwara, let L_θ be the SLD defined by Eqs. (5), (6) for the output state ρ_θ. Then the Fisher information $J(\rho_\theta)$ is given by Eq. (4). Fujiwara [13] showed that the Fisher information has a convexity property, i.e.,

$$J(\lambda\sigma_\theta + (1-\lambda)\tau_\theta) \leq \lambda J(\sigma_\theta) + (1-\lambda)J(\tau_\theta). \tag{12}$$

for any $0 < \lambda < 1$, where σ_θ and τ_θ are states with parameter θ.

To see the above relation, let Hermitian operators L_θ^σ and L_θ^τ be the SLDs of σ_θ and τ_θ, respectively, i.e.,

$$\partial_\theta \sigma_\theta = \frac{1}{2}(L_\theta^\sigma \sigma_\theta + \sigma_\theta L_\theta^\sigma), \tag{13}$$

$$\partial_\theta \tau_\theta = \frac{1}{2}(L_\theta^\tau \tau_\theta + \tau_\theta L_\theta^\tau). \tag{14}$$

Let us consider the tensor product Hilbert space $\mathcal{K} = \mathcal{H} \otimes \mathbf{C}^2$, where $\mathcal{H}$ is the state space of S and $\mathbf{C}^2$ is a 2-dimensional state space. With fixed basis $\{|0\rangle, |1\rangle\}$ of $\mathbf{C}^2$, let $\tilde\rho_\theta$ be a density operator on $\mathcal{K}$ such that

$$\tilde\rho_\theta = \lambda\sigma_\theta \otimes |0\rangle\langle 0| + (1-\lambda)\tau_\theta \otimes |1\rangle\langle 1|. \tag{15}$$

Then, it is easy to see that the SLD of $\tilde\rho_\theta$ is $L_\theta^\sigma \otimes |0\rangle\langle 0| + L^\tau \otimes |1\rangle\langle 1|$, so that the Fisher information of $\tilde\rho_\theta$ is given by

$$\begin{aligned} J(\tilde\rho_\theta) &= \mathrm{Tr}[\tilde\rho_\theta(L_\theta^\sigma \otimes |0\rangle\langle 0| + L^\tau \otimes |1\rangle\langle 1|)^2] \\ &= \lambda\mathrm{Tr}[\sigma_\theta(L_\theta^\sigma)^2] + (1-\lambda)\mathrm{Tr}[\tau_\theta(L_\theta^\tau)^2] \\ &= \lambda J(\sigma_\theta) + (1-\lambda)J(\tau_\theta). \end{aligned} \tag{16}$$

On the other hand, the partial trace of $\tilde\rho_\theta$ over $\mathbf{C}^2$ is given by

$$\mathrm{Tr}_{\mathbf{C}^2}[\tilde\rho_\theta] = \lambda\sigma_\theta + (1-\lambda)\tau_\theta. \tag{17}$$

Since the partial trace is a trace-preserving completely positive map, the monotonicity of the Fisher information under trace-preserving completely positive maps [35, 36, 37] concludes

$$J(\lambda\sigma_\theta + (1-\lambda)\tau_\theta) \leq J(\tilde\rho_\theta). \tag{18}$$

Therefore, from Eq. (16) and Eq. (18) the convexity relation (12) follows.

Now suppose that an input state $\bar\rho$ maximizes the Fisher information, i.e.,

$$J(\Gamma_\theta[\bar\rho]) = \max[J(\Gamma_\theta[\rho])]_\rho. \tag{19}$$

Let

$$\bar\rho = \sum_n p_n |n\rangle\langle n| \tag{20}$$

be the spectral decomposition, where $0 < p_n \leq 1$ and $\sum p_n = 1$. The output state $\bar\rho_\theta$ is given by

$$\bar\rho_\theta = \Gamma_\theta[\bar\rho] = \sum_n p_n \Gamma_\theta[|n\rangle\langle n|]. \tag{21}$$

By using relation (12) repeatedly, we have

$$J(\bar\rho_\theta) \leq \sum_n p_n J(\Gamma_\theta[|n\rangle\langle n|]). \tag{22}$$

Since $\bar\rho$ maximizes the Fisher information, we also have

$$\sum_n p_n J(\Gamma_\theta[|n\rangle\langle n|]) \leq J(\bar\rho_\theta), \tag{23}$$

and this concludes the relation $J(\bar\rho_\theta) = J(\Gamma_\theta[|n\rangle\langle n|])$ for all n. Thus, the maximum of the Fisher information is also attained by a pure input state.

From now on, we assume without any loss of generality that the input state of the channel is always a pure state by virtue of this theorem.

III. ONE-PARAMETER UNITARY CHANNELS HAVE NO ENHANCEMENT

Before we go to general analysis of low-noise channels, let us consider the case where the channel is unitary, or

the channel does not interact with the environment. Interestingly, the maximization of the output Fisher information $J[\rho_\theta]$ with respect to the input ρ can be explicitly accomplished. After the calculation of the maximum, one can notice that the ancilla-assisted enhancement does not take place at all. The result makes it clear that, in order to gain the ancilla-assisted enhancement for channel parameter estimations, the channels must have the effective interaction between the system and the environment.

Let $U(\theta)$ be a unitary operator with an unknown parameter θ. Then the output state of the unitary channel determined by $U(\theta)$ for an input state $|\Psi\rangle$ is given by

$$\rho(\theta) = |\Psi(\theta)\rangle\langle\Psi(\theta)|, \tag{24}$$

where the output state $|\Psi(\theta)\rangle$ is defined by

$$|\Psi(\theta)\rangle = U(\theta)|\Psi\rangle. \tag{25}$$

By introducing the (logarithmic) Hamiltonian operator $H(\theta)$ such that

$$H(\theta) = i(\partial_\theta U(\theta))U(\theta)^\dagger, \tag{26}$$
$$H(\theta)^\dagger = H(\theta), \tag{27}$$

and using the result in Ref. [5], the Fisher information of the output state is evaluated as

$$J_S[\rho_\theta] = 4V_{\Psi(\theta)}[H(\theta)], \tag{28}$$

where $V_{\Psi(\theta)}[H(\theta)]$ is the variance of $H(\theta)$ in the state $\Psi(\theta)$, i.e.,

$$V_{\Psi(\theta)}[H(\theta)] = \langle\Psi(\theta)|H(\theta)^2|\Psi(\theta)\rangle - \langle\Psi(\theta)|H(\theta)|\Psi(\theta)\rangle^2. \tag{29}$$

To obtain the maximum of J_S, let us consider the maximum and minimum of the eigenvalues E_n of $H(\theta)$:

$$E_{\max}(\theta) = \max[E_n]_n, \tag{30}$$

$$E_{\min}(\theta) = \min[E_n]_n. \tag{31}$$

Let $|\max(\theta)\rangle$ and $|\min(\theta)\rangle$ be eigenstates corresponding to $E_{\max}(\theta)$ and $E_{\min}(\theta)$, respectively. By a straightforward manipulation, it is easy to see that the maximum of J_S is taken by a pure input state $|\Phi\rangle = U(\theta)^\dagger|\Phi(\theta)\rangle$, where $|\Phi(\theta)\rangle$ is given by

$$|\Phi(\theta)\rangle = \frac{1}{\sqrt{2}}\left[|\max(\theta)\rangle + |\min(\theta)\rangle\right]. \tag{32}$$

For a fixed value of θ, the maximum is given by

$$\max[J_S]_{|\Psi\rangle} = J_S[|\Phi(\theta)\rangle\langle\Phi(\theta)|]$$
$$= (E_{\max}(\theta) - E_{\min}(\theta))^2. \tag{33}$$

To obtain the corresponding result for ancilla-assisted estimations, let us introduce an ancilla system A and the extended channel defined by

$$|\tilde{\Psi}(\theta)\rangle = (U(\theta) \otimes \mathbf{1}_A)|\tilde{\Psi}\rangle, \tag{34}$$

where $|\tilde{\Psi}\rangle$ is a state of the composite system $S + A$ to be put in the extended channel. For the output state $\tilde{\rho}(\theta) = |\tilde{\Psi}(\theta)\rangle\langle\tilde{\Psi}(\theta)|$, the Fisher information is given by

$$J_{S+A}[\tilde{\rho}(\theta)] = 4V_{\tilde{\Psi}(\theta)}[H(\theta) \otimes \mathbf{1}_A], \tag{35}$$

where $V_{\tilde{\Psi}(\theta)}[H(\theta)\otimes\mathbf{1}_A]$ is the variance of $H(\theta)\otimes\mathbf{1}_A$ in the state $\tilde{\Psi}(\theta)$. Note that the maximum and minimum of the eigenvalues of $H \otimes \mathbf{1}_A$ are taken in the states $|\max(\theta)\rangle|a\rangle$ $|\min(\theta)\rangle|a\rangle$, respectively, with an arbitrary ancilla state $|a\rangle$, i.e.,

$$H(\theta) \otimes \mathbf{1}_A|\max(\theta)\rangle|a\rangle = E_{\max}(\theta)|\max(\theta)\rangle|a\rangle, \tag{36}$$
$$H(\theta) \otimes \mathbf{1}_A|\min(\theta)\rangle|a\rangle = E_{\min}(\theta)|\min(\theta)\rangle|a\rangle. \tag{37}$$

Hence, the input state given by

$$|\tilde{\Phi}\rangle = \frac{1}{\sqrt{2}}(U(\theta)^\dagger \otimes \mathbf{1}_A)\left[|\max(\theta)\rangle|a\rangle + |\min(\theta)\rangle|a\rangle\right]$$
$$= |\Phi\rangle|a\rangle \tag{38}$$

takes the maximum value of J_{S+A}, which turns out to be the same as that given in (33), i.e.,

$$\max[J_{S+A}]_{|\tilde{\Psi}\rangle} = J_{S+A}[|\tilde{\Phi}(\theta)\rangle\langle\tilde{\Phi}(\theta)|]$$
$$= (E_{\max}(\theta) - E_{\min}(\theta))^2. \tag{39}$$

Consequently, no enhancement by the ancilla extension is observed in this unitary case, i.e.,

$$\frac{\max[J_{S+A}]_{\rho_{S+A}}}{\max[J_S]_{\rho_S}} = \frac{\max[J_{S+A}]_{|\Psi_{S+A}\rangle}}{\max[J_S]_{|\Psi_S\rangle}} = 1. \tag{40}$$

It should be noted here that the above argument applies only to one-parameter unitary channels, for which Eq. (28) can be applied, whereas a generalization to Abelian group parameters may follow. For multiple phase parameter estimation of unitary channels, Ballester [17, 18] showed, ancilla-assisted enhancement actually takes place, whereas for commuting phase parameter estimation no enhancement occurs.

Within Bayesian approach, Chiribella, D'Ariano, and Sacchi [26] showed that unitary channels with non-Abelian group parameter can have ancilla-assisted improvement of a large class of cost functions. In this connection, Sacchi [27, 28] gave extensive analysis on the condition for ancilla-assisted improvement of the error probability for discrimination of Pauli channels.

IV. LOW-NOISE CHANNELS

In this section, we introduce the notion of a low-noise channel Γ_ϵ with unknown parameter ϵ, which takes only small values $\epsilon \sim 0$, by requiring a physically natural assumption of the channel Γ_ϵ for the parameter values near $\epsilon = 0$. The small parameter ϵ is assumed to control the low noise well enough and is called the low-noise parameter.

As mentioned in the introduction, we will focus on the ancilla extension of the low-noise channel defined by $\Gamma_\epsilon \otimes id_A$. The ancilla-assisted enhancement factor η is also defined as the ratio of the Fisher information of the ancilla-assisted estimation to that of the original one and is analyzed in detail.

The concept of the noise in a quantum process to implement a target unitary process can be understood under the following consideration. Suppose that we would like to implement a unitary channel $\Lambda^{(U)}$ for a system S, so that the output state corresponding to an input state ρ_{in} of S is designed to be

$$\rho_{out} = \Lambda^{(U)}[\rho_{in}] = U\rho_{in}U^\dagger. \tag{41}$$

Without any noise, the unitary operator can be normally implemented as

$$U = \exp[-itH_S/\hbar], \tag{42}$$

where t is the time interval from input to output, and H_S is the Hamiltonian of S under control (Fig. 1).
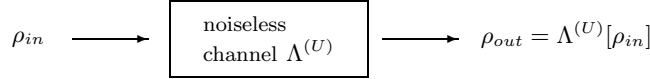

FIG. 1: Without any noise, a unitary channel $\Lambda^{(U)}$ can be implemented with a suitable controlled Hamiltonian H_S in a time interval t satisfying eq. (42).

In real life, the system S is coupled weakly with the environment E and causes the decoherence that cannot be corrected by controlling the Hamiltonian H_S of the system S, so that the noise is brought from the environment. Assume that the noise is controlled by one unknown positive parameter ν. The estimation of the noise parameter ν often becomes critical in development of quantum devices such as quantum computers.

The total Hamiltonian reads

$$H_{tot} = H_S + H_{SE} + H_E, \tag{43}$$

where H_E is the Hamiltonian of E and H_{SE} is the interaction Hamiltonian between S and E. Because of the noise, the actual output state ρ'_{out} deviates from the intended output state ρ_{out} (Fig. 2).
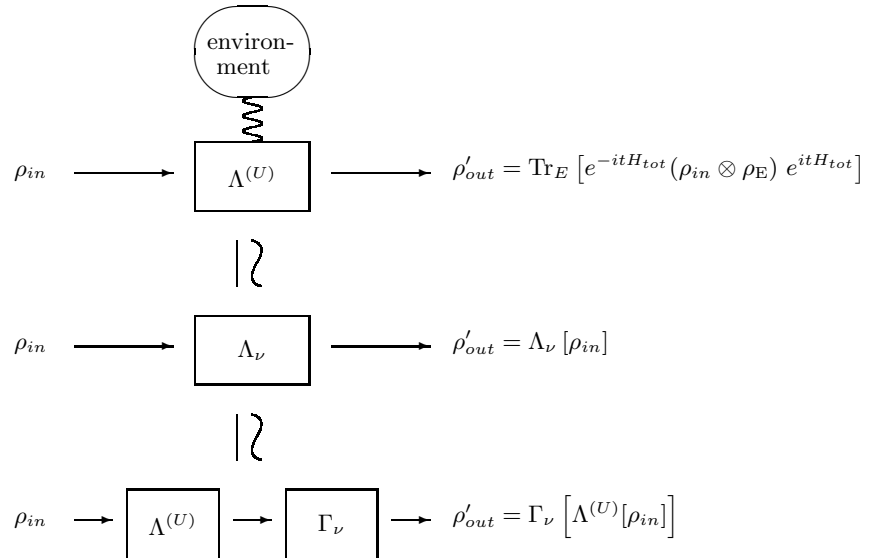

FIG. 2: A controlled unitary process, such as quantum computing, usually suffers from noise (the first line). The noisy process is described as a TPCP map, a channel Λ_ν, parametrized by one noise parameter ν (the second line). In quantum theory, the disturbed process is equivalently described by a sequence of two channels. The first channel is the originally intended unitary channel $\Lambda^{(U)}$. The second channel Γ_ν describes the genuine noise effect. We call Γ_ν the noise channel (the third line).

By using H_{tot}, the output state ρ'_{out} is determined in principle by

$$\rho'_{out} = \text{Tr}_E \left[e^{-itH_{tot}} \left(\rho_{in} \otimes \rho_E \right) e^{itH_{tot}} \right], \qquad (44)$$

where Tr_E is the partial trace over E and ρ_E is the initial state of E. Theoretically, it is preferable that we determine the value of the noise parameter ν via Eq. (44); however, the explicit calculation of Eq. (44) is too complicated to perform in many cases. Hence, adopting a reasonable theoretical model of the noise effect, the actual value of its noise parameter of the model should be experimentally estimated.

Without assuming any detailed knowledge about H_E and H_{SE}, it is natural to represent the noisy process by a TPCP map Λ_ν such that

$$\rho'_{out} = \Lambda_\nu[\rho_{in}], \qquad (45)$$

where the relation $\Lambda_0 = \Lambda^{(U)}$ holds as the noiseless case. In quantum theory, the channel Λ_ν can be equivalently described by a sequence of two channels (the third line of Fig. 2). The first one is the target unitary channel $\Lambda^{(U)}$ and the second represents the genuine noise part. This means that the general noisy process is equivalent to the noiseless unitary process followed by an instantaneous noise process. The second channel is called the noise channel Γ_ν and defined by

$$\Gamma_\nu[\rho] := \Lambda_\nu[U^\dagger \rho U] = \Lambda_\nu[(\Lambda^{(U)})^{-1}[\rho]]. \qquad (46)$$

Using the definition and the ideal output state ρ_{out}, it is possible to write the actual output state ρ'_{out} such that

$$\rho'_{out} = \Gamma_\nu[U\rho_{in}U^\dagger] = \Gamma_\nu[\rho_{out}]. \qquad (47)$$

When the noise vanishes, the channel reduces to the identity channel:
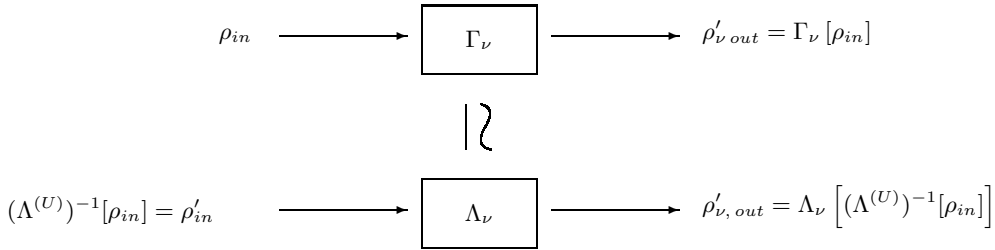
$$\Gamma_0 = id_S. \qquad (48)$$

FIG. 3: In order to experimentally generate the output state $\rho_{\nu,out}$ of the noise channel Γ_ϵ for an arbitrary input state ρ_{in}, we take $\rho'_{in} = (\Lambda^{(U)})^{-1}[\rho_{in}]$, which is independent of ν, as the input state for the actual channel Λ_ν.

It is stressed that despite that the noise channel Γ_ν is conceptual constituent, it can be simulated in a real experiment by use of the actual channel Λ_ν (Fig. 3). In fact, the output state of the channel Γ_ν defined by

$$\rho_{\nu,out} = \Gamma_\nu[\rho_{in}] \qquad (49)$$

is exactly reproduced by

$$\rho_{\nu,out} = \Lambda_\nu[(\Lambda^{(U)})^{-1}[\rho_{in}]], \qquad (50)$$

for an arbitrary input state ρ_{in}. Therefore, by adopting a known state $\rho'_{in} = (\Lambda^{(U)})^{-1}[\rho_{in}]$, which is independent of ν, as the input state of the actual channel Λ_ν, we experimentally obtain the output state $\rho_{\nu,out}$ of the noise channel Γ_ν. This aspect sounds very significant. Actually, we can replace, not only theoretically but also experimentally, the estimation problem for a given real channel Λ_ν into the equivalent estimation problem for the noise channel Γ_ν. Hence, we later concentrate on es-timation of the noise parameters for Γ_ν which satisfies relation (48).

Next let us define mathematically the low-noise channel Γ_ϵ. This is a kind of the noise channel and its noise parameter ν takes small positive values, which is denoted by ϵ. We call ϵ the low-noise parameter. Physically, Γ_ϵ is expected to have an analytic ϵ dependence near $\epsilon = 0$. A rigorous mathematical formulation of this requirement is given as follows.

Since the low-noise channel Γ_ϵ is a TPCP map, it has a Kraus representations determined by a family of Kraus operators. We shall define low-noise channels in terms of their Kraus operators. A family of TPCP maps Γ_ϵ with one parameter $\epsilon > 0$ is called a *low-noise channel with low-noise parameter* ϵ if each Γ_ϵ has a Kraus representation

$$\Gamma_\epsilon[\rho] = \sum_a B_a(\epsilon)\rho \, B_a^\dagger(\epsilon) + \epsilon \sum_\alpha C_\alpha(\epsilon)\rho \, C_\alpha^\dagger(\epsilon) \qquad (51)$$

with two classes of Kraus operators $\{B_a(\epsilon)\}$ and $\{\sqrt{\epsilon}C_\alpha(\epsilon)\}$ satisfying the following conditions:

(i) $B_a(\epsilon)$ is analytic at $\epsilon = 0$, so that we have the power series expansion

$$B_a(\epsilon) = \kappa_a \mathbf{1}_S - \sum_{n=1}^{\infty} N_a^{(n)} \epsilon^n, \qquad (52)$$

in a neighborhood of $\epsilon = 0$, where κ_a and $N_a^{(n)}$ are constant coefficients and operators, respectively, independent of ϵ. The noise channel condition in Eq. (48) requires

$$\sum |\kappa_a|^2 = 1. \qquad (53)$$

(ii) $C_\alpha(\epsilon)$ is analytic at $\epsilon = 0$, so that we have the power series expansion

$$C_\alpha(\epsilon) = M_\alpha + \sum_{n=1}^{\infty} M_\alpha^{(n)} \epsilon^n, \qquad (54)$$

in a neighborhood of $\epsilon = 0$, where M_α and $M_\alpha^{(n)}$ are constant operators independent of ϵ.

Needless to say, the Kraus operators satisfies the trace-preserving condition

$$\mathbf{1}_S = \sum_a B_a^\dagger(\epsilon)B_a(\epsilon) + \epsilon \sum_\alpha C_\alpha^\dagger(\epsilon)C_\alpha(\epsilon), \qquad (55)$$

where $\mathbf{1}_S$ is the identity operator. By definition, the relation

$$\lim_{\epsilon \to +0} \Gamma_\epsilon = id_S \qquad (56)$$

is automatically satisfied.

It should be emphasized that our definition of the low-noise channel is general from the physical point of view. Except that Γ_ϵ satisfies Eq. (56) and has analytic dependence of ϵ near the origin, the channel Γ_ϵ can be said to be a general quantum operation acting on the input state. Therefore, the low-noise channel should be always found in the weak-interaction limit of H_{SE} for rather general physical processes.

A useful comment is given here. Expanding Eq. (55) in terms of ϵ generates a lot of recursion relations between κ_a, $N_a^{(n)}$ and $M_\alpha^{(n)}$. The higher components of the operators and the coefficients are determined recursively and systematically by solving the equations using their lower components. The first-order relation in the ϵ expansion of Eq. (55) is given by

$$\sum_\alpha M_\alpha^\dagger M_\alpha = \sum_a (\kappa_a N_a^{(1)\dagger} + \kappa_a^* N_a^{(1)}). \qquad (57)$$

One of our fundamental interests is to ask a question: which input state for the low-noise channel does maximize the Fisher information of its output state ρ_ϵ? By virtue of the theorem reviewed in Section 2, the optimal input state is a pure state. Denote the input state

by $|\phi\rangle\langle\phi|$. Then, from Eq. (52) and Eq. (54), ρ_ϵ can be expanded as

$$\rho_\epsilon := \Gamma_\epsilon[|\phi\rangle\langle\phi|] = |\phi\rangle\langle\phi| - \epsilon\rho_1 + O(\epsilon^2). \qquad (58)$$

Here ρ_1 is given by

$$\begin{aligned}\rho_1 &= \sum_a [\kappa_a|\phi\rangle\langle\phi|N_a^{(1)\dagger} + N_a^{(1)}|\phi\rangle\langle\phi|\kappa_a^*] \\ &\quad - \sum_\alpha M_\alpha|\phi\rangle\langle\phi|M_\alpha^\dagger.\end{aligned} \qquad (59)$$

For this output state ρ_ϵ, we perturbatively solve the equation,

$$\partial_\epsilon \rho_\epsilon = \frac{1}{2}(L_\epsilon \rho_\epsilon + \rho_\epsilon L_\epsilon), \qquad (60)$$

in order to get the SLD operator L_ϵ. It is possible to check that the following solution actually satisfies Eq. (60) by substitution.

$$L_\epsilon = \frac{1}{\epsilon}[\mathbf{1} - |\phi\rangle\langle\phi|] - \rho_1 + O(\epsilon). \qquad (61)$$

By substituting Eq. (61) into the definition of the Fisher information, we get the value of the information such that

$$J_S[\rho_\epsilon] = \text{Tr}[\rho_\epsilon L_\epsilon^2] = \frac{1}{\epsilon}\langle\phi|\rho_1|\phi\rangle + O(\epsilon^0). \qquad (62)$$

By using Eq. (57) and Eq. (59), the Fisher information is evaluated in the leading order of ϵ as

$$J_S[\rho_\epsilon] = \frac{1}{\epsilon}\sum_\alpha \left[\langle\phi|M_\alpha^\dagger M_\alpha|\phi\rangle - |\langle\phi|M_\alpha|\phi\rangle|^2\right] + O(\epsilon^0). \qquad (63)$$

From Eq. (7) the optimal output-measurement observable A_{opt} for any input state $|\psi\rangle$ is given by $A_{opt}\rho_\epsilon = (J_S[\rho_\epsilon]^{-1}L_\epsilon + \epsilon)\rho_\epsilon$. The optimal input state $|\phi_{opt}\rangle$ can be determined by maximizing $J_S[\rho_\epsilon]$ with respect to the state $|\phi\rangle$.

Let us next discuss the low-noise channel in the ancilla-extended system $S+A$. Its extended channel is now given by $\Gamma_\epsilon \otimes id_A$. The input pure state $|\Psi\rangle$ can be decomposed into

$$|\Psi\rangle = \sum_{n=1} C_n|n\rangle \otimes |A_n\rangle, \qquad (64)$$

where $\{|n\rangle\}$ is an arbitrary orthonormal basis of S and $|A_n\rangle$'s are normalized pure states of A, which are not necessarily orthogonal to each other. The constants C_n should satisfy the normalization condition:

$$\sum_n |C_n|^2 = 1. \qquad (65)$$

The SLD for the extended state is given by

$$\tilde{L}_\epsilon = \frac{1}{\epsilon}(1 - |\Psi\rangle\langle\Psi|) + O(\epsilon^0). \qquad (66)$$

The Fisher information $J_{S+A}[\tilde{\rho}_\epsilon]$ for the output state $\tilde{\rho}_\epsilon = \Gamma_\epsilon \otimes id_A[|\Psi\rangle\langle\Psi|]$ is also evaluated in a similar manner. We have

$$J_{S+A}[\tilde{\rho}_\epsilon] = \frac{1}{\epsilon}\sum_\alpha \left[\text{Tr}[\tilde{\rho}M_\alpha^\dagger M_\alpha] - |\text{Tr}[\tilde{\rho}M_\alpha]|^2\right] + O(\epsilon^0), \tag{67}$$

where $\tilde{\rho}$ is a state of S defined by

$$\tilde{\rho} = \text{Tr}_A[|\Psi\rangle\langle\Psi|] = \sum_{n\bar{n}} C_{\bar{n}}^* \langle A_{\bar{n}}|A_n\rangle C_n |n\rangle\langle\bar{n}|. \tag{68}$$

The optimal output-measurement observable $\tilde{A}_{opt}$ for any input state $|\Psi\rangle$ is given by $\tilde{A}_{opt}\tilde{\rho}_\epsilon = (J_{S+A}[\tilde{\rho}_\epsilon]^{-1}\tilde{L}_\epsilon + \epsilon\mathbf{1})\tilde{\rho}_\epsilon$. The optimal input state $|\Psi_{opt}\rangle$ for the extended system is determined by maximizing $J_{S+A}[\tilde{\rho}_\epsilon]$ with respect to $\tilde{\rho} = \text{Tr}_A[|\Psi\rangle\langle\Psi|]$.

If the dimension of A is not less than that of S, we are able to make $|A_n\rangle$'s orthogonal to each other:

$$\langle A_{\bar{n}}|A_n\rangle = \delta_{\bar{n}n}. \tag{69}$$

Then the state $\tilde{\rho}$ is reduced into a form such that

$$\tilde{\rho} = \sum_n |C_n|^2 |n\rangle\langle n|. \tag{70}$$

Note that the orthonormal basis $\{|n\rangle\}$ of S and the coefficients C_n can be arbitrarily chosen except that $\sum|C_n|^2 = 1$. Hence, $\tilde{\rho}$ in Eq. (70) is able to describe any possible state of S. Therefore, the dimension of the ancilla Hilbert space suffices to be at most the same as the system Hilbert space.

By combining both results of J_S and J_{S+A}, we have the ancilla-assisted enhancement factor η such that

$$\eta = \frac{\max\left[\sum_\alpha \left[\text{Tr}\left[\rho_S M_\alpha^\dagger M_\alpha\right] - |\text{Tr}\left[\rho_S M_\alpha\right]|^2\right]\right]_{\rho_S}}{\max\left[\sum_\alpha \left[\langle\phi_S|M_\alpha^\dagger M_\alpha|\phi_S\rangle - |\langle\phi_S|M_\alpha|\phi_S\rangle|^2\right]\right]_{|\phi_S\rangle}}. \tag{71}$$

Here $\max[\ ]_{\rho_S}$ means the maximum value over all possible states of S and $\max[\ ]_{|\phi_S\rangle}$ the maximum value over all possible pure states of S.

Because the set of pure states of S is a subset of the set of states of S, the following inequality trivially holds:

$$\eta \geq 1. \tag{72}$$

V. EXAMPLES OF LOW-NOISE CHANNELS

Low-noise channels introduced in the previous section are found in a lot of applications. Checking that low-noise channels really appear in some physical phenomena may lead to a deeper understanding. Thus we give two critical examples in this section. The details of the channels we introduce below can be seen in Ref. [30].

A. Isotropic Depolarizing Channels

An isotropic depolarizing channel is given by

$$\Gamma_\epsilon[\rho] = \left(1 - \frac{3}{4}\epsilon\right)\rho + \frac{1}{4}\epsilon\sum_{a=1}^3 \sigma_a\rho\sigma_a. \tag{73}$$

This is a well known example induced by quantum noise. The parameter $\epsilon(\geq 0)$ is just a probability that the qubit system becomes depolarized. The Kraus operators in Eq. (52) and Eq. (54) are given by

$$B_0(\epsilon) = \left(1 - \frac{3}{4}\epsilon\right)^{1/2}\mathbf{1}_S, \tag{74}$$

$$C_a(\epsilon) = \frac{1}{2}\sigma_a, \tag{75}$$

where $a = 1, 2, 3$. Hence the expansion coefficients in Eq. (52) and Eq. (54) are given by

$$\kappa_0 = 1, \tag{76}$$

$$N_0^{(1)} = \frac{3}{8}\mathbf{1}_S, \tag{77}$$

$$N_0^{(n)} = \frac{(2n-3)!!}{n!}\left(\frac{3}{8}\right)^n \mathbf{1}_S, \tag{78}$$

$$M_a = \frac{1}{2}\sigma_a, \tag{79}$$

$$M_a^{(n)} = \mathbf{0}. \tag{80}$$

In this case, the Fisher informations have been already calculated [13]. For the isolated original system S, the information is independent of the input state and given by

$$J_S = \frac{1}{\epsilon(2-\epsilon)}. \tag{81}$$

For the extended channel $\Gamma_\epsilon \otimes id_A$, the optimal input state is the maximally entangled state and the information is given by

$$\tilde{J}_{S+A} = \frac{3}{\epsilon(4-3\epsilon)}, \tag{82}$$

as long as the parameter ϵ is small.

B. Generalized Amplitude-Damping Channels

A generalized amplitude-damping channel is given by

$$\Gamma[\rho] = \sum_{a=1}^2 B_a(\epsilon)\rho B_a^\dagger(\epsilon) + \epsilon\sum_{\alpha=1}^2 C_\alpha(\epsilon)\rho C_\alpha^\dagger(\epsilon), \tag{83}$$

where B_ν and C_α are given by

$$B_1(\epsilon) = \sqrt{\frac{1}{1+e^{-\beta E}}} \begin{bmatrix} 1 & 0 \\ 0 & \sqrt{1-\epsilon} \end{bmatrix}, \qquad (84)$$

$$B_2(\epsilon) = \sqrt{\frac{e^{-\beta E}}{1+e^{-\beta E}}} \begin{bmatrix} \sqrt{1-\epsilon} & 0 \\ 0 & 1 \end{bmatrix}, \qquad (85)$$

$$C_1(\epsilon) = \sqrt{\frac{1}{1+e^{-\beta E}}} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad (86)$$

$$C_2(\epsilon) = \sqrt{\frac{e^{-\beta E}}{1+e^{-\beta E}}} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}. \qquad (87)$$

The channel describes a relaxation process of the two-level system driven by a finite-temperature thermal bath. The temperature is $(k_B\beta)^{-1}$ where k_B is the Boltzmann constant. Here the small noise parameter ϵ is related with the survival rate s of the initial state under the relaxation such that $\epsilon = 1 - s$. The rate s is given by $s = e^{-\gamma t}$, where t is time and γ the relaxation rate constant. The corresponding coefficients in the ϵ expansion are given by

$$\kappa_1 = \sqrt{\frac{1}{1+e^{-\beta E}}}, \qquad (88)$$

$$\kappa_2 = \sqrt{\frac{e^{-\beta E}}{1+e^{-\beta E}}}, \qquad (89)$$

$$N_1^{(n)} = \sqrt{\frac{1}{1+e^{-\beta E}}} \frac{(2n-3)!!}{2^n n!} \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \qquad (90)$$

$$N_2^{(n)} = \sqrt{\frac{e^{-\beta E}}{1+e^{-\beta E}}} \frac{(2n-3)!!}{2^n n!} \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \qquad (91)$$

$$M_1 = \sqrt{\frac{1}{1+e^{-\beta E}}} \begin{bmatrix} 0 & 1 \\ 0 & 0 \end{bmatrix}, \qquad (92)$$

$$M_2 = \sqrt{\frac{e^{-\beta E}}{1+e^{-\beta E}}} \begin{bmatrix} 0 & 0 \\ 1 & 0 \end{bmatrix}, \qquad (93)$$

$$M_{1,2}^{(n)} = \mathbf{0}. \qquad (94)$$

The calculation of the Fisher information has been performed in Ref. [16].

The two examples in this section will be discussed again in Section 6.

VI. CHANNELS ON TWO-LEVEL SYSTEMS

In this section we concentrate on a two-level system S_2 and an arbitrary ancilla system A. The dimension of A is not necessarily two, but assumed finite. Let us derive a universal bound on the ancilla-assisted enhancement factor η such that

$$\eta \leq \frac{3}{2}. \qquad (95)$$

The bound must hold for all low-noise channels of S_2.

As well known, any state ρ of the two-dimensional system S_2 can be written by

$$\rho = \frac{1}{2}\mathbf{1}_S + \frac{1}{2}\vec{x} \cdot \vec{\sigma} \qquad (96)$$

where $\vec{\sigma}$ is the Pauli matrix vector and the three-dimensional real parameter vector $\vec{x}$ takes values which satisfies

$$0 \leq |\vec{x}|^2 \leq 1. \qquad (97)$$

For pure states, the vector is normal:

$$|\vec{x}|^2 = 1. \qquad (98)$$

Similarly, the matrix M_α in eqn(54) is uniquely expanded as

$$M_\alpha = m_{a0}\mathbf{1}_S + \sum_{a=1}^{3} m_{a\alpha}\sigma^a. \qquad (99)$$

Now let us define complex vectors $\vec{\mu}_a (a = 0 \sim 3)$ by using the coefficients $m_{a\alpha}$ in Eq. (99) as

$$\vec{\mu}_a = (m_{a\alpha}). \qquad (100)$$

In the vector space, there exists a natural inner product defined by

$$(\vec{u}, \vec{v}) = \sum_\alpha u_\alpha^* v_\alpha. \qquad (101)$$

A metric is also induced naturally from the inner product such that

$$g_{ab} := (\vec{\mu}_a, \vec{\mu}_b) = g_{ba}^*, \qquad (102)$$

where $a, b = 1 \sim 3$. For later convenience, define a real non-negative symmetric matrix H by

$$H = [h_{ab}] = [\text{Re } g_{ab}] \geq 0, \qquad (103)$$

and a real three-dimensional vector $\vec{J}$ by

$$\vec{J} = [J_a] = [\text{Im } g_{23}, \text{Im } g_{31}, \text{Im } g_{12}]. \qquad (104)$$

Here denote by h_1, h_2, h_3 the eigenvalues of H. Without loss of generality, we can assume that

$$0 \leq h_1 \leq h_2 \leq h_3. \qquad (105)$$

Assume later that $\vec{\mu}_a (a = 1, 2, 3)$ are linearly independent. Even if it is not so, because of the continuity of η, we can take three linearly-independent vectors $\vec{\mu}_a(t)$ parametrized by a real parameter t such that

$$\lim_{t \to 0} \vec{\mu}_a(t) = \vec{\mu}_a. \qquad (106)$$

In order to get η, we first calculate the factor $\eta(t)$ for $\{\vec{\mu}_a(t)\}$ and just take a limit as

$$\lim_{t \to 0} \eta(t) = \eta. \qquad (107)$$

Note that the linearly independence of $\{\vec{\mu}_a\}$ also means

$$H > 0. \qquad (108)$$

This allows us to assume the existence of H^{-1}.

By a simple manipulation, we have

$$\eta(\vec{J}) = \frac{\mathrm{Tr}H + \vec{J}H^{-1}\vec{J} - \min\left[(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})\right]_{|\vec{x}|\leq 1}}{\mathrm{Tr}H + \vec{J}H^{-1}\vec{J} - \min\left[(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})\right]_{|\vec{x}|=1}}. \tag{109}$$

For the original system S, the optimal input state is given by

$$|\phi\rangle\langle\phi| = \frac{1}{2}\mathbf{1}_S + \frac{1}{2}\vec{x}_{opt}\cdot\vec{\sigma} \tag{110}$$

where $\vec{x}_{opt}$ is the vector which minimizes $(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})$ among whole the unit vectors. The optimal input state $|\Psi\rangle$ for the extended system is also given as follows. Find a vector $\vec{X}$ which minimizes $(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})$ among whole the vectors with $|\vec{x}| \leq 1$. Then the optimal state $|\Psi\rangle$ is determined by solving the equation

$$Tr_A[|\Psi\rangle\langle\Psi|] = \frac{1}{2}\mathbf{1}_S + \frac{1}{2}\vec{X}\cdot\vec{\sigma}. \tag{111}$$

Let us consider the case where $\vec{J} = \vec{0}$. The factor η is given by

$$\begin{aligned}
\eta &= \frac{\mathrm{Tr}H - \min\left[\sum_{a,b=1}^{3} x^a x^b h_{ab}\right]_{|\vec{x}|\leq 1}}{\mathrm{Tr}H - \min\left[\sum_{a,b=1}^{3} x^a x^b h_{ab}\right]_{|\vec{x}|=1}} \\
&= \frac{h_1 + h_2 + h_3}{h_1 + h_2 + h_3 - \min\left[\sum_{a=1}^{3} h_a(x'^a)^2\right]_{|\vec{x}'|=1}}.
\end{aligned}$$

Here we have made H diagonalized in the last equality. Consequently we obtain an expression of η such that

$$\eta = \frac{h_1 + h_2 + h_3}{h_1 + h_2 + h_3 - \min[h_1, h_2, h_3]}. \tag{112}$$

Taking account of $h_1 \leq h_2 \leq h_3$, we can easily prove $\eta \leq 3/2$ as follows.

$$\begin{aligned}
\eta &= \frac{h_1 + h_2 + h_3}{h_2 + h_3} \\
&\leq \frac{2h_2 + h_3}{h_2 + h_3} \\
&\leq \frac{3h_3}{2h_3} = \frac{3}{2}.
\end{aligned} \tag{113}$$

Next let us discuss the case where $\vec{J} \neq \vec{0}$. Suppose that $|H^{-1}\vec{J}| \leq 1$. Then we have

$$\min\left[(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})\right]_{|\vec{x}|\leq 1} = 0, \tag{114}$$

because we can always take a vector $\vec{x}$ such that $\vec{x} = -H^{-1}\vec{J}$. For later convenience, let us introduce a function $G(\vec{J})$ as

$$G(\vec{J}) := \min\left[(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})\right]_{|\vec{x}|=1}. \tag{115}$$

Then we can prove that the function G satisfies

$$G(\vec{0}) \geq G(\vec{J}).$$

To show this, we transform $G(\vec{J})$ as

$$G(\vec{J}) = \min\left[\vec{X}H\vec{X}\right]_{|\vec{X}-H^{-1}\vec{J}|=1}. \tag{116}$$

By denoting $\vec{K} = H^{-1}\vec{J}$, the function G is given in the diagonal basis of H by

$$G = \min\left[\sum_a h_a(X'_a)^2\right]_{|\vec{X}'-\vec{K}'|=1}. \tag{117}$$

Note that the relation $|\vec{K}'| \leq 1$ trivially holds. Also notice from definition (115) that if $\vec{J} = \vec{K}' = \vec{0}$, G takes the minimum value of the eigenvalues of H, that is, h_1:

$$G(\vec{0}) = h_1. \tag{118}$$

To compare $G(\vec{J})$ with this value h_1, suppose a point $\vec{X}'_o$ on a trajectory defined by $|\vec{X}' - \vec{K}'| = 1$ such that

$$\vec{X}'_o = \left(K'_1 \pm \sqrt{1 - (K'_2)^2 - (K'_3)^2}, 0, 0\right). \tag{119}$$

Then we have

$$\vec{X}'_o H \vec{X}'_o = h_1\left(K'_1 \pm \sqrt{1 - (K'_2)^2 - (K'_3)^2}\right)^2. \tag{120}$$

Here we fix the double sign in the above equation so as to satisfy the relation:

$$\vec{X}'_o H \vec{X}'_o = h_1\left(|K'_1| - \sqrt{1 - (K'_2)^2 - (K'_3)^2}\right)^2. \tag{121}$$

Since the relation $|\vec{K}'| \leq 1$ holds, it is guaranteed that

$$1 \geq \left(|K'_1| - \sqrt{1 - (K'_2)^2 - (K'_3)^2}\right)^2. \tag{122}$$

Therefore the important inequality

$$G(\vec{0}) \geq G(\vec{J}) \tag{123}$$

really arises as follows.

$$\begin{aligned}
G(\vec{0}) &= h_1 \\
&\geq h_1\left(|K'_1| - \sqrt{1 - (K'_2)^2 - (K'_3)^2}\right)^2 = \vec{X}'_o H \vec{X}'_o \\
&\geq \min\left[\vec{X}'H\vec{X}'\right]_{|\vec{X}'-\vec{K}'|=1} = G(\vec{J}).
\end{aligned} \tag{124}$$

Note that

$$\vec{J}H^{-1}\vec{J} \geq 0 \qquad (125)$$

and

$$\mathrm{Tr}H - G(\vec{0}) = h_2 + h_3 > 0. \qquad (126)$$

Keeping Eq. (123), Eq. (125) and Eq. (126) in mind, let us go back to the proof of $\eta(\vec{J}) \leq \eta(\vec{0})$. By using Eq. (114), we have

$$\eta(\vec{J}) = \frac{\mathrm{Tr}H + \vec{J}H^{-1}\vec{J}}{\mathrm{Tr}H + \vec{J}H^{-1}\vec{J} - G(\vec{J})}. \qquad (127)$$

By replacing $G(\vec{J})$ by $G(\vec{0})$ in the above equality, from Eq. (123) we obtain

$$\eta(\vec{J}) \leq \frac{\mathrm{Tr}H + \vec{J}H^{-1}\vec{J}}{[\mathrm{Tr}H - G(\vec{0})] + \vec{J}H^{-1}\vec{J}}. \qquad (128)$$

By using an inequality such that

$$\frac{b+\epsilon}{a+\epsilon} \leq \frac{b}{a} \qquad (129)$$

for $a \leq b$ and $\epsilon \geq 0$ with $\epsilon = \vec{J}H^{-1}\vec{J}$, we have

$$\eta(\vec{J}) \leq \frac{\mathrm{Tr}H}{\mathrm{Tr}H - G(\vec{0})} = \eta(\vec{0}). \qquad (130)$$

Consequently, we have obtained the bound

$$\eta(\vec{J}) \leq \eta(\vec{0}) \leq 3/2. \qquad (131)$$

For the remaining case where $|H^{-1}\vec{J}| > 1$, the problem becomes much trivial. This is because

$$\min\left[(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})\right]_{|\vec{x}| \leq 1}$$
$$= \min\left[(\vec{x} + H^{-1}\vec{J})H(\vec{x} + H^{-1}\vec{J})\right]_{|\vec{x}| = 1} \quad (132)$$

holds in this case. Therefore the relation $\eta = 1$ is satisfied in Eq. (109).

Therefore, for all the possible low-noise channels, the bound $\eta \leq 3/2$ has been proven. The equality $\eta = 3/2$ can be attained by the channels satisfying

$$g_{ab} \propto \delta_{ab} \qquad (133)$$

with the maximally-entangled input pure states of $S + A$.

The optimal input state depends on the vector $\vec{J}$ of the channel. When $\vec{J} = \vec{0}$, the optimal input state is the maximally entangled state. If $|H^{-1}\vec{J}| \geq 1$, a factorized input state takes the maximum and gives $\eta = 1$. When $1 > |H^{-1}\vec{J}| > 0$, the optimal input state is neither the maximally entangled state nor the factorized state. From the argument below Eq. (114) the output state $|\psi\rangle_{S+A}$ satisfies

$$\mathrm{Tr}_A[|\psi\rangle_{S+A}\langle\psi|_{S+A}] = \frac{1}{2}\mathbf{1}_S - \frac{1}{2}\vec{J}H^{-1}\vec{\sigma}. \qquad (134)$$

The value of η given by Eq. (127) also changes continuously between $1 \leq \eta < 3/2$ depending on $|H^{-1}\vec{J}|$.

The channel dependence of the optimal input state has been already noticed in a generalized amplitude-damping channel [16] by changing the temperature of the thermal bath. Because of the simplicity of the model, it is possible to estimate the unknown parameter even in a finite parameter region. On the other hand, in this paper, the parameter region of the low-noise channel is constrained to a neighborhood of a fixed value ($\epsilon = 0$). However, we would like to stress that our channel includes an enormous number of degrees of freedom corresponding to κ_a, $N_a^{(n)}$ and $M_\alpha^{(n)}$, compared with the generalized amplitude-damping channel.

Note that the isotropic depolarizing channel (Eq. (73) in Section 5) is one of the channels attaining the bound ($\eta = 3/2$). The vectors $\vec{\mu}_a$ are calculated as

$$\vec{\mu}_1 = \frac{1}{2}\begin{bmatrix} 1 \\ 0 \\ 0 \end{bmatrix}, \qquad (135)$$

$$\vec{\mu}_2 = \frac{1}{2}\begin{bmatrix} 0 \\ 1 \\ 0 \end{bmatrix}, \qquad (136)$$

$$\vec{\mu}_3 = \frac{1}{2}\begin{bmatrix} 0 \\ 0 \\ 1 \end{bmatrix}. \qquad (137)$$

The corresponding matrix g_{ab} is just evaluated as

$$g_{ab} = \frac{1}{4}\delta_{ab}. \qquad (138)$$

Thus the channel can achieve $\eta = 3/2$.

On the other hand, the generalized amplitude-damping channels Eq. (83) in Section 5) cannot achieve the bound. The vectors $\vec{\mu}_a$ are now described by

$$\vec{\mu}_1 = \frac{1}{2}\begin{bmatrix} \sqrt{\frac{1}{1+e^{-\beta E}}} \\ \sqrt{\frac{e^{-\beta E}}{1+e^{-\beta E}}} \\ 0 \end{bmatrix}, \qquad (139)$$

$$\vec{\mu}_2 = \frac{i}{2}\begin{bmatrix} \sqrt{\frac{1}{1+e^{-\beta E}}} \\ -\sqrt{\frac{e^{-\beta E}}{1+e^{-\beta E}}} \\ 0 \end{bmatrix}, \qquad (140)$$

$$\vec{\mu}_3 = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \qquad (141)$$

The corresponding g_{ab} is now given by

$$[g_{ab}] = \frac{1}{4}\begin{bmatrix} 1 & i\frac{1-e^{-\beta E}}{1+e^{-\beta E}} & 0 \\ -i\frac{1-e^{-\beta E}}{1+e^{-\beta E}} & 1 & 0 \\ 0 & 0 & 0 \end{bmatrix}. \qquad (142)$$

Because $g_{ab} \propto \delta_{ab}$ does hold, the channel cannot satisfy $\eta = 3/2$ for any parameter value. In spite of the ancilla

extension, the ancilla-assisted enhancement does not appear at all ($\eta = 1$), as long as the low-noise parameter ϵ is small enough. This is because the value of $|H^{-1}\vec{J}|$ diverges and the relation $|H^{-1}\vec{J}| > 1$ always holds.

[1] C. W. Helstrom, *Quantum Detection and Estimation Theory* (Academic, New York, 1976).
[2] A. S. Holevo, *Probabilistic and Statistical Aspects of Quantum Theory* (North-Holland, Amsterdam, 1982).
[3] M. Hayashi, ed., *Asymptotic theory of quantum statistical inference: Selected papers* (World Scientific, Singapore, 2005).
[4] M. Hotta and M. Ozawa, Phys. Rev. A **70**, 022327 (2004).
[5] A. Fujiwara and H. Nagaoka, Phys. Lett. A **201**, 119 (1995).
[6] I. L. Chuang and M. A. Nielsen, J. Mod. Opt. **44**, 2455 (1997).
[7] J. F. Poyatos, J. I. Cirac, and P. Zoller, Phys. Rev. Lett. **78**, 390 (1997).
[8] K. Kraus, Ann. Phys. (N.Y.) **64**, 311 (1971).
[9] K. Kraus, *States, Effects, and Operations: Fundamental Notions of Quantum Theory*, Lecture Notes in Phys. **190** (Springer, Berlin, 1983).
[10] M. Ozawa, in *Probability Theory and Mathematical Statistics,* Lecture Notes in Math. **1021**, edited by K. Itô and J. V. Prohorov (Springer, Berlin, 1983), pp. 518–525.
[11] M. Ozawa, J. Math. Phys. **25**, 79 (1984).
[12] M. Ozawa, Ann. Phys. (N.Y.) **311**, 350 (2004).
[13] A. Fujiwara, Phys. Rev. A **63**, 042304 (2001).
[14] A. Fujiwara, Phys. Rev. A **65**, 012316 (2002).
[15] A. Fujiwara and H. Imai, J. Phys. A: Math. Gen. **36**, 8093 (2003).
[16] A. Fujiwara, Phys. Rev. A **70**, 012317 (2004).
[17] M. A. Ballester, Phys. Rev. A **69**, 022303 (2004).
[18] M. A. Ballester, Phys. Rev. A **70**, 032310 (2004).
[19] A. Fujiwara, in *Quantum Information, Statistics, Probability*, edited by O. Hirota (Rinton Press, Princeton, 2004), pp. 78–89.
[20] D. G. Fischer, H. Mack, M. A. Cirone, and M. Freyberger, Phys. Rev. A **64**, 022309 (2001).
[21] T. C. Bschorr, D. G. Fischer, and M. Freyberger, Phys. Lett. A **292**, 15 (2001).
[22] M. Sasaki, M. Ban, and S. M. Barnett, Phys. Rev. A **66**, 022308 (2002).
[23] A. M. Childs, J. Preskill, and J. Renes, J. Mod. Opt. **47**, 155 (2000).
[24] A. Acín, Phys. Rev. Lett. **87**, 177901 (2001).
[25] G. M. D'Ariano, P. Lo Presti, and M. G. A. Paris, Phys. Rev. Lett. **87**, 270404 (2001).
[26] G. Chiribella, G. M. D'Ariano, and M. F. Sacchi, *Optimal estimation of group transformations using entanglement* (2005), online preprint: quant-ph/0506267.
[27] M. F. Sacchi, Phys. Rev. A **71**, 062340 (2005).
[28] M. F. Sacchi, *Minimum error discrimination of pauli channels* (2005), online preprint: quant-ph/0506072.
[29] G. M. D'Ariano, M. F. Sacchi, and J. Kahn, *Minimax discrimination of two pauli channels* (2005), online preprint: quant-ph/0507081.
[30] M. A. Nielsen and I. L. Chuang, *Quantum Computation and Quantum Information* (Cambridge University Press, Cambridge, 2000).
[31] P. H. Eberhard, Tech. Rep. 72-1, CERN (1972).
[32] W. C. Carithers, J. H. Christenson, P. H. Eberhard, D. R. Nygren, T. Modis, T. P. Pun, E. L. Schwartz, and H. Sticker, Phys. Rev. D **14**, 290 (1976).
[33] J. Ellis, J. S. Hagelin, D. V. Nanopoulos, and M. Srednicki, Nucl. Phys. B **241**, 381 (1984).
[34] P. Huet and M. E. Peskin., Nucl. Phys. B **434**, 3 (1995).
[35] D. Petz, Linear Algebr. Appl. **244**, 81 (1996).
[36] E. A. Morozova and N. N. Centsov, Itogi Nauki Tehniki **36**, 69 (1990).
[37] D. Petz and C. Sudár, J. Math. Phys. **37**, 2662 (1996).